A Nanometer-Level Pathlength Control Scheme for Nulling Interferometry

E. Serabyn

Nulling interferometry[1-3] is a promising technique for reducing a star's brightness relative to its immediate surroundings, with great potential to enable direct detections of extra-solar planets and zodiacal light. The technique is based on the precise cancellation, or "nulling", of the starlight received by two separate telescopes, and so the amplitudes, phases and polarizations of the two on-axis electric fields must all be matched to very high accuracy across the waveband of interest. A workable realization for a nulling beam combiner (NBC) is then far from trivial. The issue addressed herein is that of achieving the necessary degree of pathlength control in an NBC: starlight cancellation at the 10^{-6} level in the mid-infrared (MIR), as envisaged for NASA's planned Terrestrial Planet Finder[4] (TPF), or equivalently, at the 10^{-4} level at visible wavelengths, as planned for NASA's Space Interferometer Mission[4], requires pathlength control and stabilization at the nanometer level, an order of magnitude finer than typical interferometric requirements. Here, an optical control scheme which can achieve the requisite degree of pathlength control is presented.

To cancel on-axis starlight to high accuracy, the electric fields from two telescopes viewing a common star must be combined exactly out of phase at all wavelengths across the band of interest. One clever method for introducing an achromatic π-radian phase flip between two beams is a geometric flip of the electric field vector, such as is provided by a rotational shearing interferometer[5-7]. Two different implementations of this idea have been proposed, one of which uses orthogonal rooftop reflectors in the two arms of an interferometric beam combiner (in which each rooftop rotates the output electric field vector by 90 deg in opposite senses), while the second uses cat's eye reflectors in an interferometric beam combiner (and relies on passage through focus in one of the NBC arms but not the other). As the rooftop version is the more symmetric, the proposed NBC control scheme is developed in that context here. However, it applies equally well to all rotational shearing interferometers which have two identical "balanced" outputs (see below).

The issue of NBC pathlength control at the nanometer level is not simply one of precision, but first and foremost one of defining a viable control scheme. The reason is quite simple: the goal of nulling a star's light calls for reducing the signal photons by a factor as large as 10^4 to 10^6, thus seemingly reducing the signal available for pathlength control by the same factor. Furthermore, ignoring transmission leakage sources other than the phase difference, ϕ, between the two combining beams, at any wavelength the transmission of a NBC is given by

$$T = \frac{P_{\text{out}}}{P_{\text{in}}} = \frac{1}{2}(1 - \cos\phi), \tag{1}$$

where P_{in} is the single aperture input power and P_{out} is the power emerging in a single output. (The negative sign arises from the π-radian phase flip). For small ϕ, i.e. near the null position,

$$\frac{P_{\text{out}}}{P_{\text{in}}} = (\frac{\phi}{2})^2. \tag{2}$$

Two further disadvantages to use of the transmitted, or "leakage" power can be discerned in this relationship: at null ($\phi = 0$), the transmission vs. phase relationship has zero slope, and the leakage power is independent of the sign of ϕ. The leakage power thus does not provide an unambiguous off-null error signal without the introduction of some type of modulation (the second harmonic of a sinusoidal phase error would contain the necessary information, albeit at low signal to noise). Of course out-of-band, and hence non-nulled, radiation may also be capable of serving as an off-null error signal, but such an approach remains to be developed.

To overcome these limitations, a new control scheme has been devised which is based upon a unique and heretofore unrecognized property of interferometric beam combiners. Before describing the concept in detail, it is first necesary to further define the term "interferometric beam combiner", or IBC, as used here. Two individual input beams from a pair of telescopes can of course be superposed upon each other at a simple, single-pass beamsplitter, thus converting the two telescopes into a single-baseline interferometer. In this case, the "beam combiner" is simply a single optical element. In both outputs emerging from the beamsplitter (one to either of its sides), one of the input beams is modulated by the beamsplitter's electric field transmission coefficient, t, and the other by its reflection coefficient, r. Pathlength matching is accomplished prior to the beamsplitter via active "delay line" optical elements, and the position of zero relative delay between the beams is typically found by a fringe search technique.

In contrast, in an IBC, the combining element is itself a small interferometer, in which a beamsplitter is both the input and output element, and the two input beams are superposed upon each other at the second beamsplitter pass as a result of lateral beam shifts induced by the mirrors located in each arm of the IBC (Fig. 1). The beamsplitter is thus used in double pass, and as a result of the lateral beam shifts, four non-equivalent outputs are produced. Two outputs are "balanced", in the sense that in each of them, the two input fields are modulated by the same beamsplitter rt product (with a relative phase difference of π radians introduced by the action of the orthogonal rooftops). On the other hand, the two additional outputs which are returned toward the input side are unbalanced: in each of these, one input field is modulated by r^2 and the other by t^2. Furthermore, in these unbalanced outputs, the π phase difference is negated (as must be the case for energy conservation to hold) by the unbalanced traversals of the beamsplitter. Such a beam combiner is thus an "interferometer within an interferometer", and so is referred to as an IBC here. An IBC is certainly more complex than a beam combiner consisting of a single beamsplitter, but an IBC is preferable in certain contexts (such as nulling) both because of the achromatic π phase shift which is automatically introduced between the two IBC arms, and also because of the automatic power matching between the two inputs (both of which are modulated by the same rt product) which takes place in each of the balanced outputs (thus removing the need to perfectly match r against t, as in the single beamsplitter case).

In an IBC, the existence of two balanced and completely equivalent outputs leads to the possibility of complete phase control. The basic idea is to control one of the balanced outputs by means of the second, thereby allowing active stabilization of the first output at the null position, at the cost of half the input power. All that then remains is to define the mechanism. It turns out that this can be accomplished quite simply, by making use of a rather unique property of an IBC: if a path offset, x_{int}, is introduced into one arm of an IBC, the phases of the two balanced outputs are affected in opposite senses. In terms of the two input beams, one of the balanced outputs emerges with beam 1 advanced by x_{int} relative to beam 2, while simultaneously in the other balanced output, beam 2 is advanced by x_{int} relative to beam 1 (Fig. 1). Thus for internal (i.e., inside the IBC) path errors around the null position, the two balanced outputs depart from null in opposite directions.

This property of IBCs can be exploited if an external path delay, x_{ext}, is also inserted into one of the input beams prior to the IBC. The total relative delay between the two input beams seen in the two balanced outputs will then be $x_{ext} - x_{int}$ and $x_{ext} + x_{int}$, respectively. If now x_{ext} is set equal to x_{int}, the relative path delay in the first balanced output becomes exactly zero (i.e., this output is nulled), while the second balanced output is offset relative to the null position by $2x_{int}$ (Fig. 2). Note that the nulling output remains achromatically zeroed while the offset output acquires chromatic properties; however, only the nulled output needs to be achromatic anyway.

Control of the null is enabled by the offset output, for which the output power can vary much more rapidly with phase than near null. The off-null error signal can be maximized by setting $2x_{int} = \lambda_{\rm m}/4$, where $\lambda_{\rm m}$ is the mean wavelength of the radiation, because this offset maximizes the slope of the transmission vs. phase relation (Eq. 1). At this "quadrature" operating point, the output power is at half-maximum, and the fractional power change vs. the phase error ϕ is

$$\frac{\Delta P}{P} = \phi = 2\pi \frac{\Delta x}{\lambda_{\rm m}}, \tag{3}$$

where Δx is the corresponding pathlength error. At $\lambda_{\rm m} \approx 630$ nm, the sensitivity is then a healthy 1% per nanometer (while at $\lambda = 12$ μm, the sensitivity is still 0.05%/nm). Source-intrinsic fluctuations (e.g., in a laboratory demonstration) can be held well below the 1% level by normalizing the quadrature-offset output by the full-power non-nulling, or constructive, output (the complement of the nulling output), implying that sub-nanometer accuracy can be attained.

In comparison, at the nulling output a 1 nm path error induces an increase in output power at $\lambda = 630$ nm of only 2.5×10^{-5} of the input power, while in the MIR case of $\lambda = 12$ μm, this factor drops to a meager 2.5×10^{-7} (here however, shorter wavelengths offer an alternative control route). While in principle such small power changes should nevertheless be detectable, obviously the quadrature output provides a much higher signal-to-noise ratio (SNR) than dithering about the null position can. This translates directly into a much higher control bandwidth, as to achieve a given rms photon shot noise (or SNR), the necessary integration time is inversely proportional to the photon flux. The control bandwidth thus scales linearly with the photon flux, and hence is several orders of magnitude larger when using the quadrature output. This gain is particularly important in low flux conditions, such as small collecting apertures and long wavelengths.

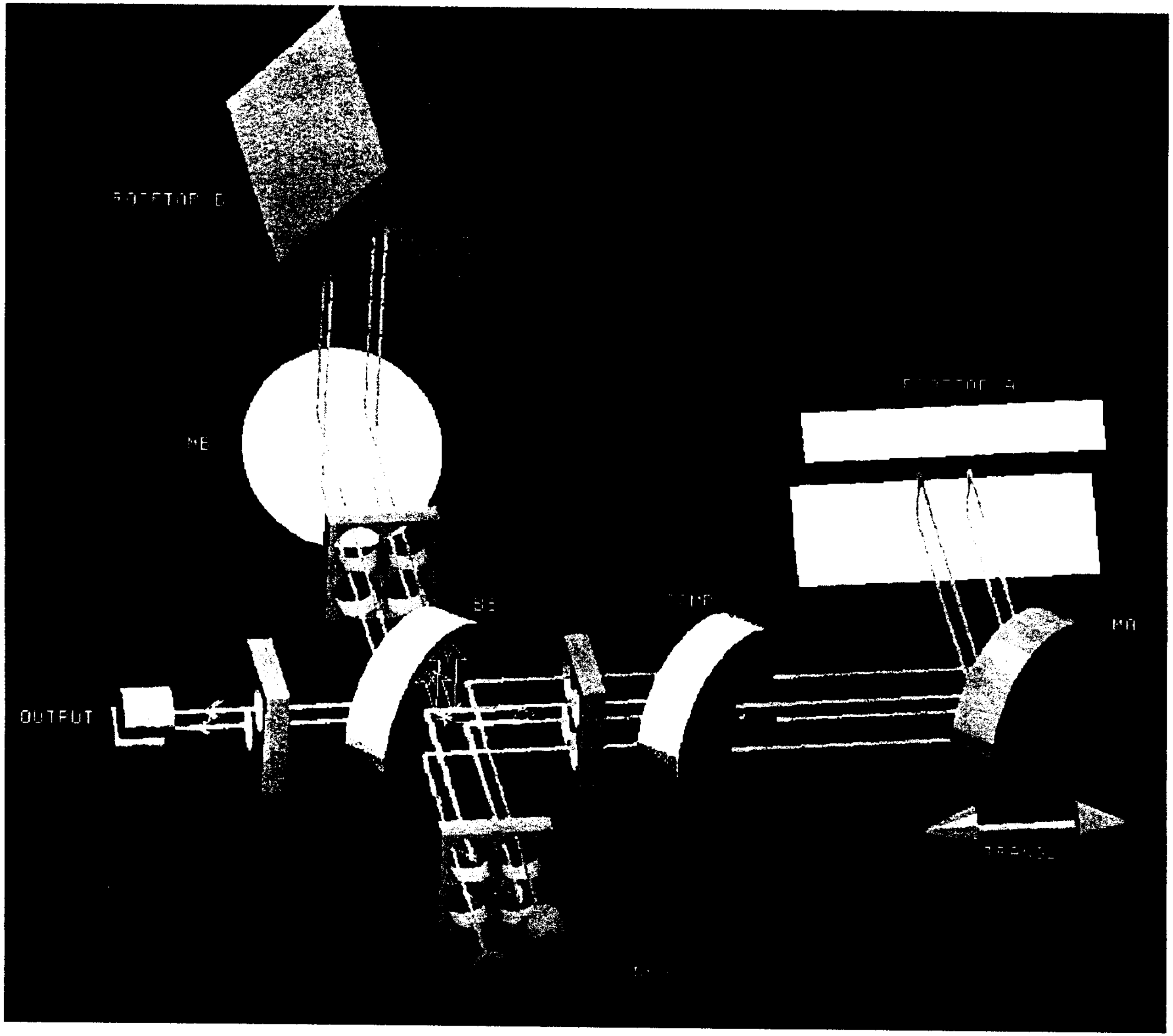
To complete the sensing scheme, only one further point need be addressed, that being how to distinguish between variations of x_{int} and x_{ext}. To enable this discrimination, laser metrology can be used to monitor x_{int} variations inside the NBC itself. In an NBC based on a pair of crossed rooftop reflectors behind a beamsplitter, the most direct approach for standard heterodyne laser metrology would be to inject the laser beam down the axis defined by the virtual intersection of the crossed rooftop axes (Fig. 3), as then the return beams from the rooftops would retrace their paths. If necessary, small corner cubes can be attached to the rooftops at these locations to remove tilts in the reflected wavefronts. As the NBC beamsplitter is not expected to be of the polarizing type, the two NBC arms would also require orthogonally-oriented plane polarizers inserted at the corner cubes in order to restrict the signals traveling down the two NBC arms to single, orthogonal polarizations, as called for by standard heterodyne metrology implementations.

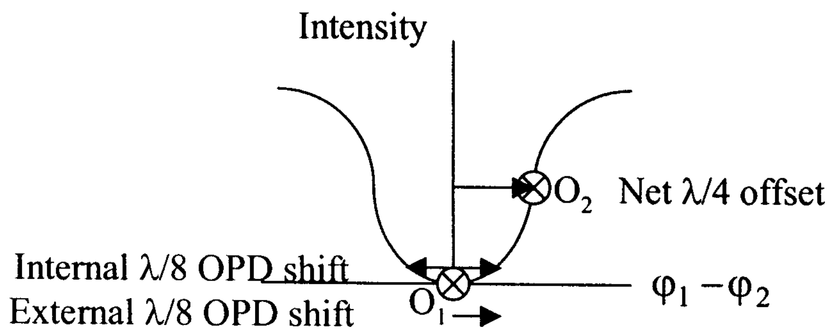
Assuming that NBC-internal metrology is implemented, and that a means of introducing an external pathlength delay is present (such as the delay lines present in stellar interferometers), all of the signals necessary for complete control of an NBC at the nanometer level are then in place. Once a null-fringe search is carried out, and the two balanced outputs are set at the null and quadrature locations, the combination of the NBC-internal metrology and the (normalized) power variations at the quadrature output can be used to track these locations. The metrology signal gives x_{int} variations directly, while the difference between the two signals gives x_{ext} variations. Using such a quadrature output, it should thus be possible to maintain a very deep null for extended periods.

Acknowledgements

References

Dual beam rooftop nuller



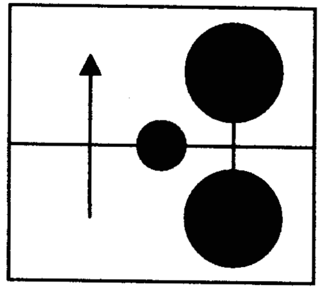

Intensity
O_2 Net λ/4 offset
Internal λ/8 OPD shift
External λ/8 OPD shift
O_1
$\varphi_1 - \varphi_2$

2

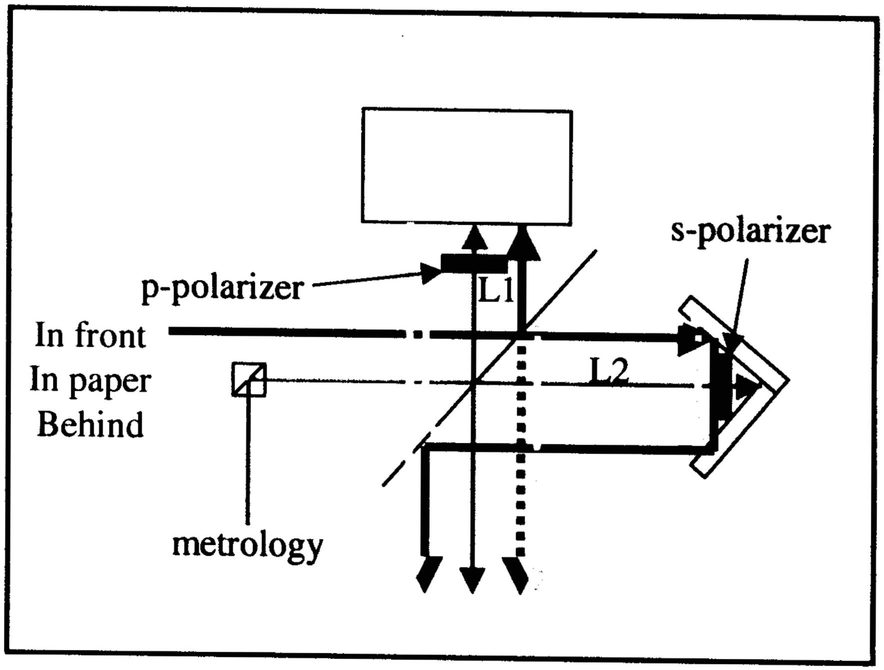
10

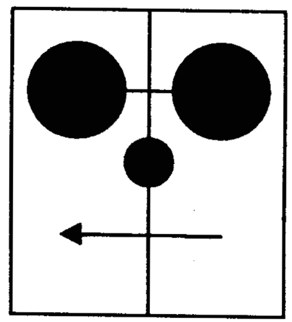
Metrology: along rooftop axis

- On-axis metrology gives internal (L1-L2) OPD variations



(Views looking back at rooftop faces from output beams)





W